IMAGE GATHERING AND CODING FOR DIGITAL RESTORATION: INFORMATION EFFICIENCY AND VISUAL QUALITY

Friedrich O. Huck
NASA Langley Research Center, Hampton, Virginia

Sarah John, Judith A. McCormick, and
Ramkumar Narayanswamy
Science and Technology Corporation, Hampton, Virginia

ABSTRACT

Image gathering and coding are commonly treated as tasks separate from each other and from the digital processing used to restore and enhance the images. Our goal in this paper is to develop a method that allows us to assess quantitatively the combined performance of image gathering and coding for the digital restoration of images with high visual quality. Digital restoration is often interactive because visual quality depends on perceptual rather than mathematical considerations, and these considerations vary with the target, the application, and the observer. Our approach is based on the theoretical treatment of image gathering as a communication channel [J. Opt. Soc. Am. A2, 1644 (1985); 5,285 (1988)]. Initial results suggest that the practical upper limit of the information contained in the acquired image data ranges typically from ~ 2 to 4 binary information units (bifs) per sample, depending on the design of the image-gathering system. The associated information efficiency of the transmitted data (i.e., the ratio of information over data) ranges typically from ~ 0.3 to 0.5 bif per bit without coding to ~ 0.5 to 0.9 bif per bit with lossless predictive compression and Huffman coding. These upper limits of performance are reached when the sampling passband of the image-gathering system closely matches the Wiener spectrum of the incident radiance field. The visual quality that can be attained with interactive image restoration improves perceptibly as the available information increases to ~ 3 bifs per sample. However, the perceptual improvements that can be attained with further increases in information are very subtle and depend on the target and the desired enhancement.

1. INTRODUCTION

Image gathering and coding are commonly treated as tasks separate from each other and from the digital processing used to restore and enhance the images. Ordinarily, image-gathering systems are designed to produce good visual quality for conventional image displays, and data-compression techniques are developed to reduce, as much as possible, the data necessary to reproduce a faithful duplicate of this original image. Image restoration and enhancement, despite the rapidly increasing use of digital processing, are virtually ignored in these assessments of image gathering and coding.

Digital image restoration is often interactive because a single figure of merit for visual quality does not exist to formulate a single "best" algorithm. Visual quality is too elusive a concept for such a figure to exist. It depends on a number of attributes, such as fidelity (resemblance to the scene), resolution (minimum discernible detail), sharpness (contrast between large areas), and clarity (absence of visual artifacts and noise). The trade-off between these attributes of image quality still must be based on perceptual rather than mathematical considerations, and these considerations vary with the target, the application, and the observer. In addition, sometimes the enhancement of certain target features is desirable to improve resolution and sharpness, even at the cost of fidelity and clarity.

In previous papers Huck et al.[1,2] have shown that image gathering can be treated like a communication channel if (and only if) the image-gathering degradations are correctly accounted for in image processing. If this is done, then the informationally optimized image-gathering system tends to maximize the fidelity and robustness of a variety of optimally restored representations ranging from images to edges. It also is possible with interactive image restoration to improve significantly on the visual quality produced by the traditional methods employed in digital image gathering and restoration.[3,4] These traditional methods[5-10] often have failed to improve on the visual quality obtained in a simpler and faster way by image reconstruction and interpolation. It is perhaps for this reason, at least in part, that image gathering and coding have not been assessed directly with digital restoration and enhancement in the past.

In this paper we extend the information theoretic assessment of image gathering to include image coding. Our goal is to develop a method that allows us to assess quantitatively the combined performance of image gathering and coding for the interactive restoration of images with high visual quality. The major questions we deal with are: How much visual information can be acquired by the image-gathering process? How much information is required to restore images with high visual quality? And how can this information be transmitted most efficiently? We do not attempt here to compare the performance of a variety of image-coding techniques. Instead, we limit our assessment to the familiar lossless predictive compression with Huffman coding.

2. OUTLINE

Figure 1 presents the end-to-end block diagram of the image gathering, coding, and restoration processes that we analyze in this paper. Our approach is to assess quantitatively the flow of information through the image gathering and coding processes, followed by a qualitative assessment of the visual quality that can be restored from the transmitted information.

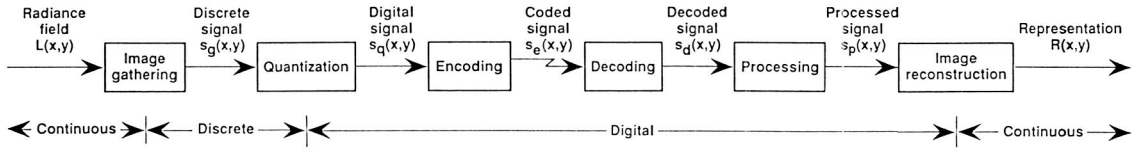


Figure 1. Model of image gathering, coding, and restoration.

In Section 3 we assess the information density of the data that is acquired by the image-gathering process in terms of the Wiener spectrum of the radiance field, the design of the image- gathering system, and the dynamic range and quantization intervals of the quantizer. We also introduce the concept of information efficiency (i.e., the ratio of information over data) as an additional criterion of the effectiveness of image gathering and coding. Our formulations are based on the theoretical treatment of image gathering given by Huck et al.[1,2] Following the methods of Shannon[11] and of Fellgett and Linfoot,[12] this treatment is constrained by the assumption that the radiance field and the noise are wide-sense-stationary Gaussian random processes.

In Section 4 we assess the effects of image coding on the information efficiency of the transmitted data. We use the familiar lossless predictive compression together with Huffman coding. The predictive compression reduces the statistical redundancy of the digital data without loss of information, and the Huffman coding compresses the data by transmitting the more probable symbols in fewer bits than the less probable ones. In addition, we demonstrate that important differences exist between the information density of the transmitted data and the entropy that is often used in the prevailing digital processing literature[5-10] to assess data compression.

In Section 5 we assess the quality of the restored images as a function of the available information. We first consider fidelity-maximized restorations. These restorations allow us to perform parametric trade-offs in terms of a single figure of merit, namely, the image fidelity. However, the fidelity-maximized images exhibit some visual defects such as ringing, aliasing artifacts, and noise. Thus, we use the Wiener-Gaussian enhancement (WIGE) filter introduced by McCormick et al.[4] to suppress these defects and improve the visual quality.

3. IMAGE GATHERING

A. Information Capacity

Let the image-gathering system acquire information about some isoplanatism area A of the radiance field $L(x, y)$ with the average power σ_L^2. Furthermore, let the image-gathering process be constrained, like a communication channel, only by the frequency passband $\widehat{B}$ and the white noise $n(x, y)$ with the power σ_N^2. Then the absolute upper limit of the acquired information about the area A is defined by the expression

$$H_c = \frac{1}{2}|A||\widehat{B}\log_2\left[1 + (K\sigma_L/\sigma_N)^2\right],$$

where $K\sigma_L/\sigma_N$ is the rms signal-to-noise ratio (SNR) and K is the steady-state gain of the radiance-to-signal conversion in the image-gathering process. The associated information capacity of the image-gathering process per unit area A and unit passband $\widehat{B}$ is

$$h_c = \frac{H_c}{|A||\widehat{B}|} = \frac{1}{2}\log_2\left[1 + (K\sigma_L/\sigma_N)^2\right]. \tag{1}$$

The magnitude of h_c may be defined as bifs, binary information units per unit area and passband, analogous to bits for the binary units per sample of the transmitted digital data.

The information capacity h_c is plotted as a function of the SNR $K\sigma_L/\sigma_N$ in Fig. 2. It varies from ~ 1 bif for $K\sigma_L/\sigma_N = 2$ to ~ 10 bifs for $K\sigma_L/\sigma_N = 1000$. This is the range of SNR's that is typically of interest. SNR's below this range ordinarily do not permit the restoration of images with good visual quality, and SNR's above this range ordinarily do not improve the visual quality. In practice, of course, information is inevitably lost because the Wiener spectrum of random radiance fields is not white and band-limited to $\widehat{B}$ and because the image-gathering process introduces aliasing and blurring as well as noise. In addition, information ordinarily is lost by the signal quantization that is required for digital data transmission and processing.

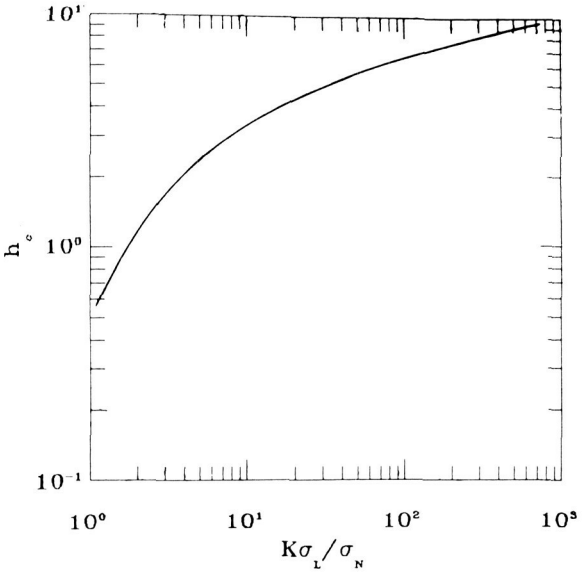


Figure 2. Information capacity h_c versus SNR $K\sigma_L/\sigma_N$. The image-gathering process is assumed to be constrained, like a communication channel, only by the sampling passband $\widehat{B}$ and the SNR $K\sigma_L/\sigma_N$. The radiance field spectrum is assumed to be white and band limited to $\widehat{B}$.

B. Radiance-Field Properties

We assume that the incident radiance field $L(x, y)$ consists of contiguous rectangles whose sides are parallel to some axes (x', y') (see Fig. 3). The transitions along each axis obey the Poisson probability-density function with the (expected) mean separation λ^{-1}, and the radiance-field magnitude of each rectangle obeys the zero-mean Gaussian probability-density function with the (expected) variance σ_L^2. The resultant autocorrelation of $L(x, y)$ is[13]

$$\begin{aligned}\Phi_L(x, y) &= \sigma_L^2 \exp\left[-\lambda(1-c)\left(|x'| + |y'|\right)\right] \\ &= \sigma_L^2 \exp\left[-\left(|x'| + |y'|\right)/\mu\right]\end{aligned}$$

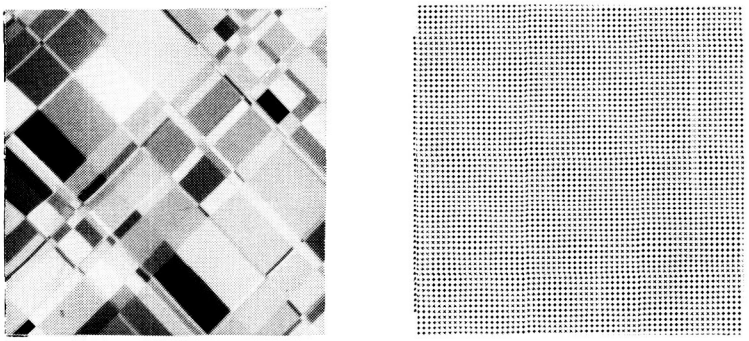
where c is the correlation of the radiance-field magnitudes of adjacent rectangles and, for convenience, $\mu = 1/\lambda(1-c)$. If we let the orientation of the (x', y') axes of the rectangles be random with uniform probability, then the autocorrelation becomes circularly symmetric as given by

$$\Phi_L(x, y) = \Phi_L(r) = \sigma_L^2 \exp\left(-|r|/\mu\right), \tag{2a}$$

where $r^2 = x^2 + y^2$. The corresponding Wiener spectrum of $L(x, y)$ can be closely approximated by[1-4,14,15]

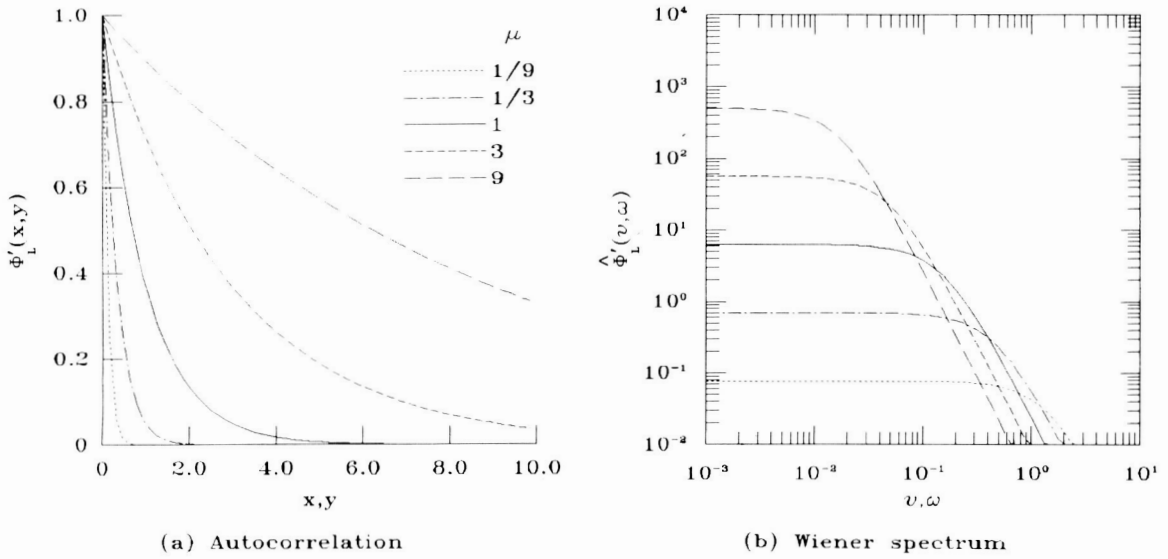
$$\widehat{\Phi}_L(v, w) = \widehat{\Phi}_L(\rho) = \frac{2\pi\mu^2\sigma_L^2}{\left[1 + (2\pi\mu\rho)^2\right]^{3/2}}, \tag{2b}$$

where $\rho^2 = v^2 + w^2$. Figure 4 illustrates the normalized auto correlation $\Phi'_L(x, y) = \sigma_L^{-2}\Phi_L(x, y)$ and the normalized Wiener spectrum $\widehat{\Phi}'_L(v, w) = \sigma_L^{-2}\Phi_L(v, w)$. The mean spatial detail of the radiance field is conveniently represented by μ. This implies that the correlation is $c \approx 0.3$. The exact expression for the Wiener spectrum of the target shown in Fig. 3 is given by Fales et al.[3] [Eq. (18)]. The Wiener-spectrum curves for that expression are almost identical to those given for the more convenient Eq. (2b).



(a) Target **(b) Sampling lattice**

Figure 3. Random radiance field with mean spatial detail $\mu = 3$, and the sampling lattice with unit sampling intervals.



(a) Autocorrelation (b) Wiener spectrum

Figure 4. Autocorrelation functions and Wiener spectra of the radiance field for several mean spatial details μ.

C. Image-Gathering Degradations

Conventional image-gathering systems consist of an objective lens (or lens system) and some sort of photon-detection and sampling mechanism. The most common mechanisms are sensor-array and line-scan devices. The lens and photosensor apertures are basically low-pass spatial-frequency filters. The spatial-frequency response of the image-gathering system, which is the product of these two low-pass-filter responses, ordinarily decreases smoothly with increasing spatial frequency until the lens diffraction limit is reached.

Figure 5 presents a model of the image-gathering process that transforms the continuous radiance field $L(x, y)$ into the signal $s_g(x, y)$ as defined by the expression

$$s_g(x, y) = [KL(x, y) * \tau_g(x, y)]\, \underline{\underline{|||}}(x, y) + n(x, y), \tag{3a}$$

where K is the steady-state gain of the (linear) radiance-to-signal conversion, $n(x, y)$ is the (additive, discrete) sensor noise, $*$ denotes convolution, and $\underline{\underline{|||}}(x, y)$ denotes sampling. Taking the Fourier transform of $s_g(x, y)$ yields the spatial-frequency representation of the acquired signal

$$\hat{s}_g(v, w) = K\hat{L}(v, w)\hat{\tau}_g(v, w) * \hat{\underline{\underline{|||}}}(v, w) + \hat{n}(v, w), \tag{3b}$$

where $\hat{L}(v, w)$ and $\hat{n}(v, w)$ are the spatial-radiance and noise transforms, respectively, and $\hat{\tau}_g(v, w)$ is the spatial-frequency response of the image-gathering system. The sampling function is given by

$$\begin{aligned} \hat{\underline{\underline{|||}}}(v, w) &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(v - m, w - n) \\ &= \delta(v, w) + \underset{\neq 0,0}{\hat{\underline{\underline{|||}}}}(v, w) \end{aligned}$$

for unit sampling intervals. The term $\underset{\neq 0,0}{\hat{\underline{\underline{|||}}}}(v, w)$ accounts for the sampling sidebands.

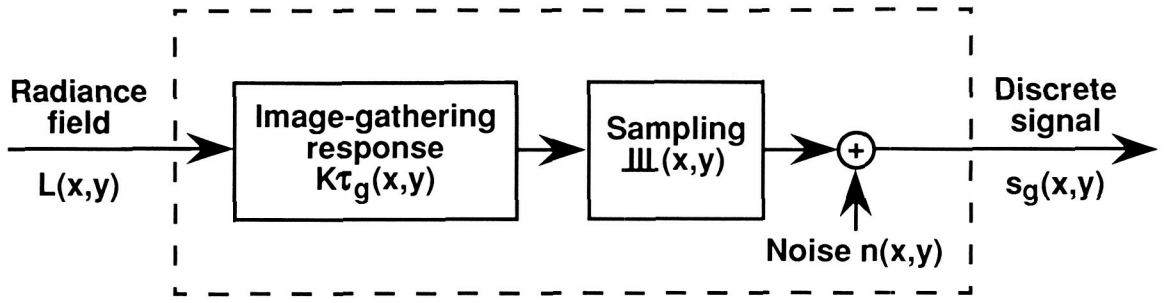


Figure 5. Model of the image-gathering process.

It is convenient to normalize the sampling intervals to unity, and to define the frequency passband $\widehat{B}$ as the sampling passband given by

$$\widehat{B} = \{(v, w), |v| < 0.5, \ |w| < 0.5\}. \tag{4}$$

The corresponding area in the frequency domain is $|\widehat{B}| = 1$. The low-pass frequency response of conventional image-gathering systems can often be approximated by the Gaussian form[16]

$$\hat{\tau}_g(v, w) = \exp\left[-(\rho/\rho_c)^2\right], \tag{5}$$

where the optical-design parameter ρ_c is the spatial frequency at which $\hat{\tau}_g(v, w) = 1/e = 0.37$.

If we now let the image-gathering process be constrained by the response $\hat{\tau}_g(v, w)$, the sampling passband $\widehat{B}$, and the SNR $K\sigma_L/\sigma_N$, then the information density h_g of the acquired signal $s_g(x, y)$ becomes[1]

$$h_g = \frac{1}{\widehat{B}} \iint \log_2 \left[1 + \frac{\widehat{\Phi}'_L(v, w)|\hat{\tau}_g(v, w)|^2}{\Phi'_L(v, w)|\hat{\tau}(v, w)|^2 * \widehat{\coprod}_{\neq 0, 0}(v, w) + (K\sigma_L/\sigma_N)^{-2}} \right] dv \ dw. \tag{6}$$

The associated variance σ_g^2 of the signal is

$$\sigma_g^2 = \sigma_L^2 \int_{-\infty}^{\infty} \widehat{\Phi}'_L(v, w)|\hat{\tau}_g(v, w)|^2 \ dv \ dw. \tag{7}$$

Figures 6 and 7 illustrate the dependence of the information density h_g on the optical-design parameter p_c for several mean spatial details μ and SNR's $K\sigma_L/\sigma_N$. These results suggest the two following generalizations:

(1) The information density h_g tends to be maximum when the sampling passband most closely matches the radiance-field spectrum, regardless of the design of the image-gathering system. This occurs, for the target characterized by Eqs. (2), when the sampling intervals are approximately equal to the mean spatial detail (i.e., when $\mu \approx 1$).

This generalization intuitively is appealing when one considers image restoration. One could not expect to restore spatial detail that is much finer than the sampling interval, and one ordinarily could expect to restore detail that is much coarser from fewer samples.

(2) The informationally optimized optical design (i.e., trade-off between aliasing and blurring) is a function of the SNR.

Again, this generalization intuitively is appealing when one considers image restoration. In one extreme, when the SNR is very low, then the restoration of fine detail is constrained by noise, and so it ordinarily would be preferable to avoid substantial blurring (at the cost of aliasing). In the other extreme, when the SNR is very high, then the restoration of fine detail is not constrained by noise, and so it ordinarily would be preferable to avoid substantial aliasing (at the cost of blurring). However, as we will show in Section 5 below, some other constraints may be introduced by the image restoration, such as ringing near sharp edges (Gibb's phenomenon).

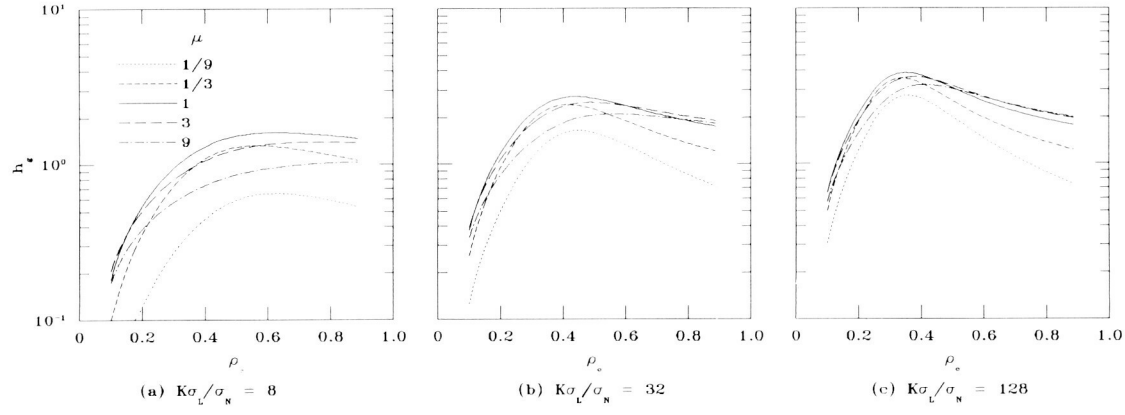


Figure 6. Variation of information density with image-gathering system design. Results are given for several mean spatial details μ and SNR's $K\sigma_L/\sigma_N$.

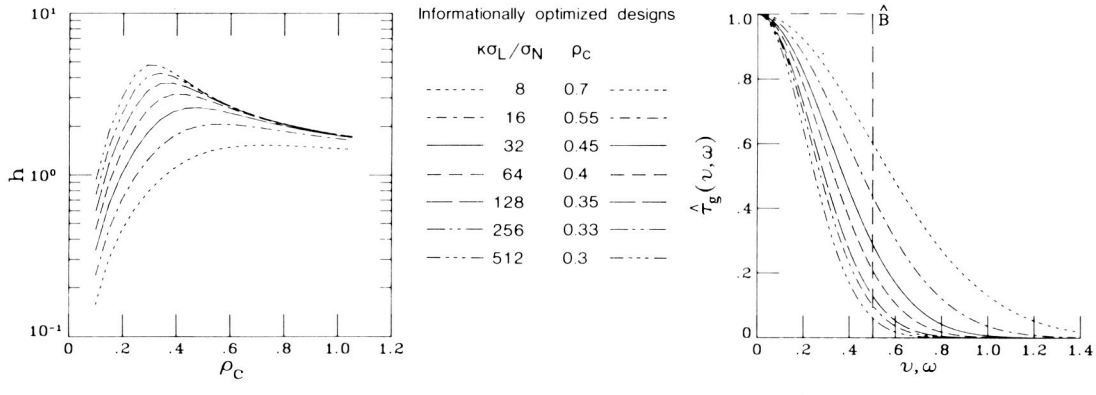


Figure 7. Variation of information density with image-gathering system design. The informationally optimized design is the design for which the image-gathering response designated by ρ_c is selected to maximize the information density h_g for a given SNR $K\sigma_L/\sigma_N$.

As a consequence of the above two generalizations, we limit the following quantitative investigations mostly to the mean spatial detail $\mu = 1$ but consider three optical designs (see Fig. 7) throughout the remainder of this paper. They are (a) the conventional response $\rho_c = 0.7$ that is also informationally optimum for very low SNR's, (b) the response $\rho_c = 0.45$ that is informationally optimum for intermediate SNR's, and (c) the response $\rho_c = 0.35$ that is informationally optimum for very high SNR's.

D. Quantization

Each discrete signal $s_g(x, y)$ is quantized into κ levels for η-bit encoding, $\kappa = 2^\eta$. If we divide the area A into M by N samples, then the area of A for unit sampling intervals is $|A| = MN$. Thus, the number of distinguishable states in A is κ^{MN}, and the amount of data in A is

$$H_d = MN \log_2 \kappa.$$

The associated data density is

$$h_d = \frac{H_d}{MN} \log_2 \kappa = \eta. \tag{8}$$

It is convenient to let the units of h_d be bits even though strictly they are bits per sample. Just as the information capacity h_c given by Eq. (1) sets a theoretical upper limit on the information density h_g acquired by image gathering, so the data density h_d given by Eq. (8) sets a theoretical upper limit on the information density transmitted by digital communication.

E. Information Efficiency

Since it ordinarily is desirable to use as few encoding levels as possible, some loss of information density is associated with the quantization process. Hence, the information density h_q of the quantized signal $s_q(x, y)$ is closely interrelated with the data density h_d. This interrelationship suggests the definition of information efficiency as the ratio h_q/h_d. The units of this ratio are bif/bit. This definition of information efficiency is analogous to Khinchin's definition of "relative entropy" as the ratio $h/\log m$, where h is the entropy of the test, and $\log m$ is the maximum value of h for the m different symbols of the test.[17] Another analogy is Jones's definition of "information efficiency" of a light beam as the information capacity per transmitted photon.[18]

To properly interpret the information efficiency h_q/h_d, we must account for an important difference between continuous and discrete entropies. The data density h_d is defined for a *discrete* random variable (i.e., the quantization levels with a uniform probability density function) for which the entropy provides i an absolute measure of randomness. The information density h_q is defined for a ***continuous*** random variable (i.e., the continuous magnitude with a Gaussian probability density function) for which the entropy provides a measure of randomness relative to an assumed standard. It intuitively is satisfying to adjust the ratio h_q/h_d so that the theoretical upper limit of information efficiency becomes unity. This adjustment occurs when (1) the Wiener spectrum $\widehat{\Phi}_L(v, w)$ is white and band-limited to $\widehat{B}$, (2) the image-gathering response $\hat{\tau}_g(v, w)$ is unity within $\widehat{B}$ and zero outside, and (3) the quantization intervals are very large compared with the magnitude of the noise. The information density h_q of the digital data becomes then

$$h_q = \frac{1}{2} \iint_{\widehat{B}} \log_2 \left[1 + \frac{\widehat{\Phi}'_L(v, w) |\hat{\tau}_g(v, w)|^2}{\widehat{\Phi}'_L(v, w) |\hat{\tau}_g(v, w)^2 * \underset{\neq 0,0}{\widehat{\text{III}}}(v, w) + (K\sigma_L/\sigma_N)^{-2} + \kappa^{-2}} \right] dv\, dw. \tag{9}$$

Equation (9) for the information density h_q reduces to Eq. (8) for the data density h_d when the above three conditions are evoked. The final step in this reduction of Eq. (9) to Eq. (8) entails the approximation given by

$$h_q = \frac{1}{2} \log_2(1 + \kappa^2) \approx \log_2 \kappa = h_d.$$

This adjustment of the information density h_q leads to a linear encoding of the Gaussian signal variation over a dynamic range of $-\sqrt{3}K\sigma_g$ to $\sqrt{3}K\sigma_g$, which encompasses 92% of the signal. The corresponding quantization interval is $\Delta = 2\sqrt{3}K\sigma_g/\kappa$. Values of $s_g(x,y) < \bar{s}_g - \sqrt{3}\sigma_g$ are assigned the value 0 and values of $s_g(x,y) > \bar{s}_g + \sqrt{3}\sigma_g$ are assigned the value $\kappa - 1$, where $\bar{s}_g$ is the average value of $s_g(x,y)$.

Figures 8 and 9 illustrate the dependence of the information density h_q and the information efficiency h_q/h_d on the optical-design parameter ρ_c, the SNR $K\sigma_L/\sigma_N$, and the number of encoding levels η. These results suggest the following generalizations:

(3a) Conventional optical responses ($\rho_c = 0.7$) limit the information density to $h_q \approx 2.2$ bifs. This limit is closely approached when the SNR is $K\sigma_L/\sigma_N \approx 20$ and the number of encoding levels is $\eta \approx 6$ bits. The corresponding maximum information efficiency is $h_q/h_d \approx 0.53$ bif/bit, but with the reduced information density $h_q \approx 1.7$ bifs as obtained with 3-bit encoding.

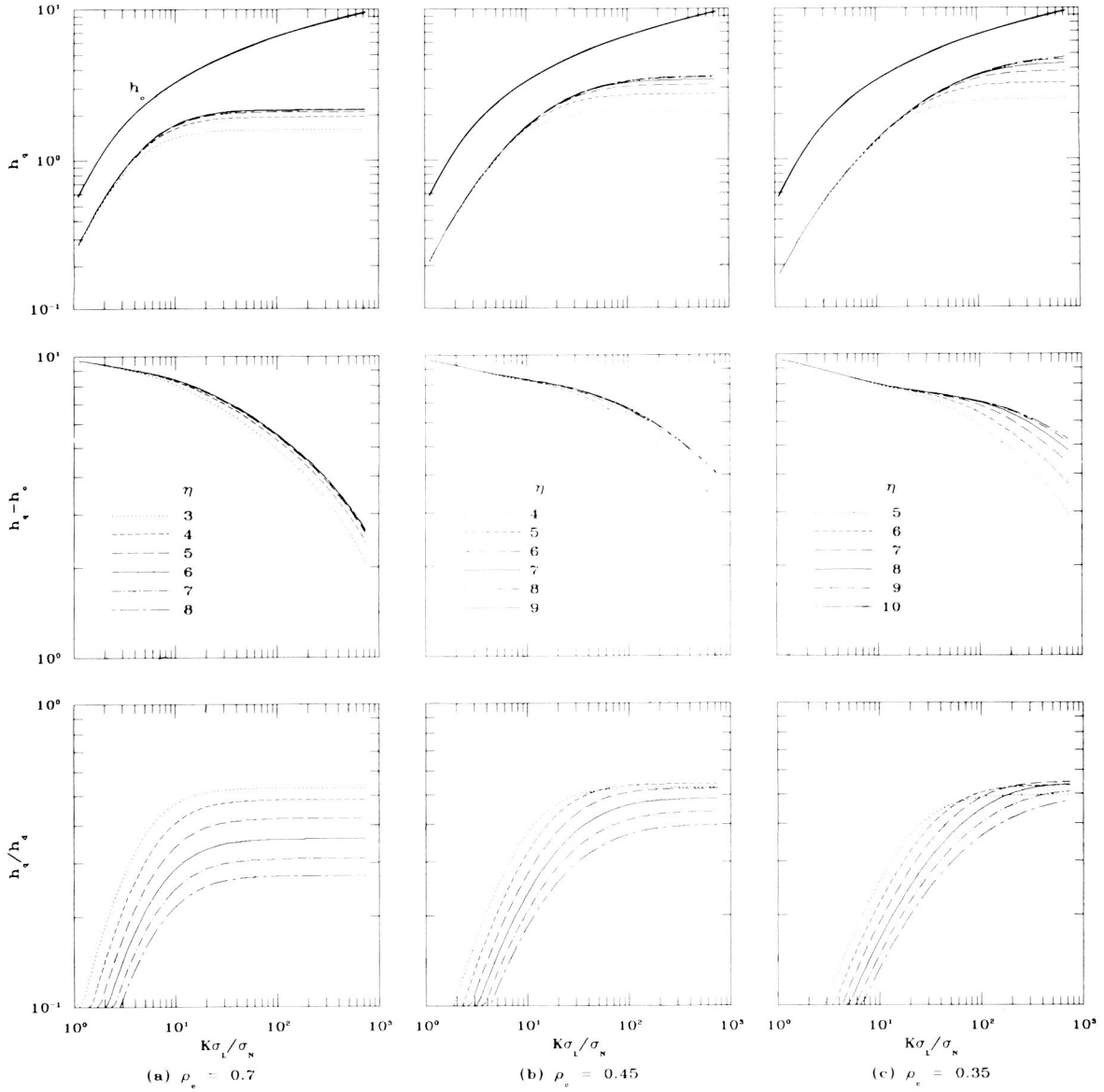


Figure 8. Information versus the SNR $K\sigma_L/\sigma_N$ for several encoding levels η. The first row presents the information capacity h_c and density h_q, the second row presents the lost information $h_q - h_c$, and the third row presents the information efficiency h_q/h_d. The image-gathering system is characterized by the optical-design parameter ρ_c and the SNR $K\sigma_L/\sigma_N$. The mean spatial detail $\mu = 1$.

(3b)Optical responses ($\rho_c = 0.45$) that are informationally optimized for intermediate SNR's limit the information density to $h_q \approx 3.6$ bifs. This limit is closely approached when the SNR is $K\sigma_L/\sigma_N$ 80 and the number of encoding levels is $\eta \approx 7$. The corresponding maximum information efficiency is $h_q/h_d \approx 0.54$ bif/bit, but with the reduced information density $h_q \approx 2.7$ bifs as obtained with 5-bit encoding.

(3c)Optical responses ($\rho_c = 0.35$) that are informationally optimized for high SNR's limit the information density to $h_q \approx 4.7$ bifs. This limit is closely approached when the SNR is $K\sigma_L/\sigma_N \approx 240$ and the number of encoding levels is $\eta \approx 8$. The corresponding maximum information efficiency is $h_q/h_d \approx 0.55$ bif/bit, but with the reduced information density $hq \approx 3.7$ bifs as obtained with 7-bit encoding.

The preferred number of encoding levels for information density, ranging from $\eta \approx 6$ for low SNR's to $\eta \approx 8$ for high SNR's, corresponds closely to those often encountered in practice. However, their selection entails some conflict between information density and efficiency. This conflict resolves itself with data compression. Furthermore, as these results foreshadow, it is the image-gathering system that is designed for highest information density that also provides the highest information efficiency with lossless image coding.

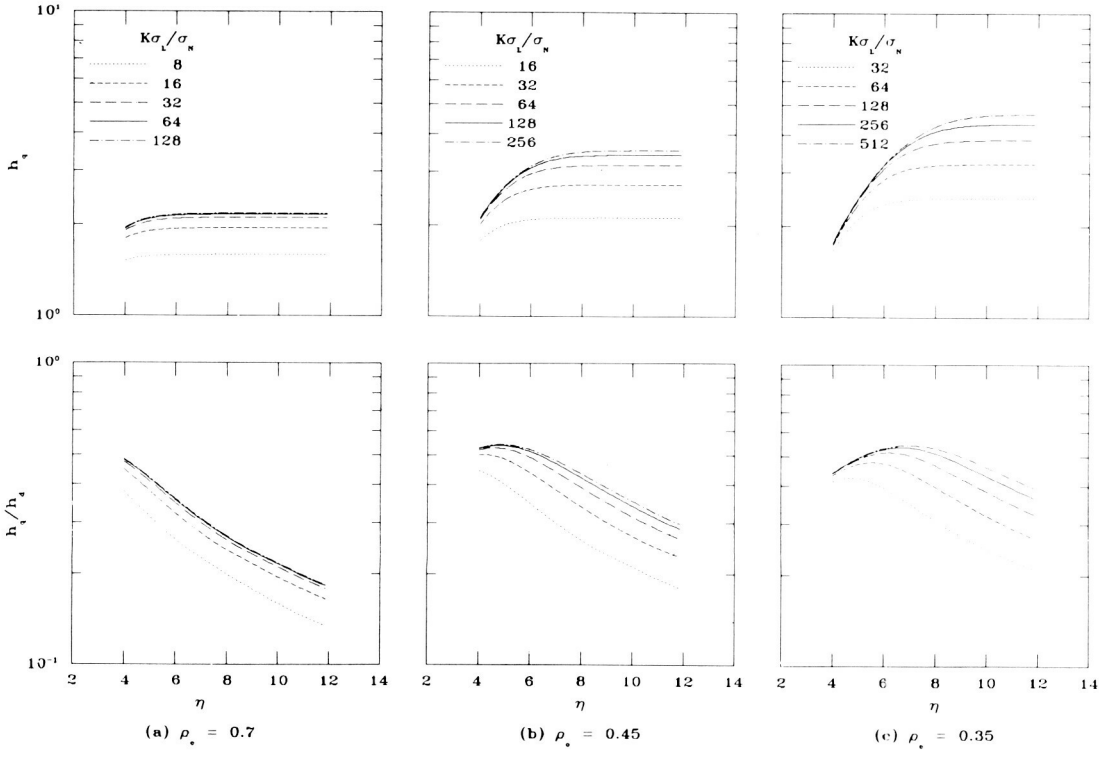


Figure 9. Information density h_q and efficiency h_q/h_d versus the encoding level η for several SNR's. The image-gathering system is characterized by the optical-design parameter ρ_c and the SNR $K\sigma_L/\sigma_N$. The mean spatial detail $\mu = 1$.

4. DATA COMPRESSION

A. Entropy Versus Information Density

It is common in the prevailing literature[5-10] to consider information density to be synonymous with the entropy used to assess data compression. However, entropy, unlike information density, does not distinguish between the properties of the scene that we wish to restore and the degradations of the image-gathering process that we wish to minimize. For example, whereas aliasing and noise subtract from the information density, these same degradations add to the entropy. Since it is not possible to distinguish quantitatively between the desired and undesired components of the signal, it also is not possible to measure the information density of an image. However, it is possible, at least in theory, to measure its entropy.

The entropy of the digital signal $s_q(x, y)$ can be determined as follows. Let $p(x_1, x_2, ..., x_{MN})$ be the probability of realizing a particular set of digital data containing MN samples and κ quantization levels per sample. Then the entropy θ_q of this data is defined as the logarithm of the probable number of alternate, distinguishable sets given by

$$\theta_q = -\frac{1}{MN} \sum_{x_1=1}^{\kappa} \sum_{x_2=1}^{\kappa} \cdots \sum_{x_{MN}=1}^{\kappa} p(x_1, x_2, \ldots, x_{MN}) \log_2 p(x_1, x_2, \ldots, x_{MN}). \tag{10}$$

The entropy θ_q given by Eq. (10) is equal to the information density h_g given by Eq. (9) only if the undesired components of the signal are negligible. However, since these undesired components are ordinarily not negligible, the information density h_g seldom reaches the entropy θ_q (i.e., $h_q < \theta_q$).

The amount of computation required to find θ_q given by Eq. (10) is, in practice, prohibitive. An upper boundary for the entropy θ_q can readily be found if we assume that each sample is independent of its neighbors, i.e., that

$$p(x_1, x_2, \ldots, x_{MN}) = p(x_1)p(x_2) \ldots p(x_{MN}),$$

where $x_i = \Delta k_i$ and k_i is an integer, $1 \leq k_1 \leq \kappa$ [Fig. 10(a)]. Letting $p_i \equiv p(x_i) \equiv p(\Delta k_i)$, the upper boundary for θ_q is then given by

$$\theta_{q0} = -\sum_{i=1}^{\kappa} p_i \log_2 p_i. \tag{11}$$

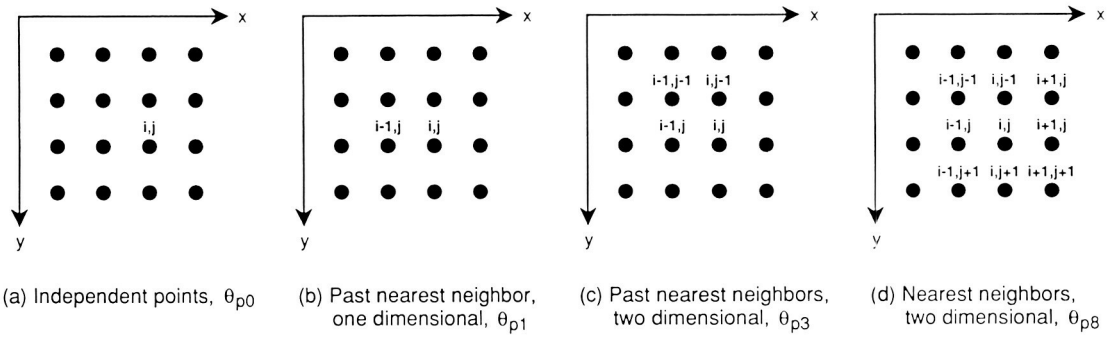
Ordinarily, the values p_i are obtained from the probability distribution (histogram) of the digital data $s_q(x, y)$. If we were to assume that the probability distribution is uniform so that all quantization levels are equally likely, then $p_i = 1/\kappa$ and the upper boundary θ_{q0} of the entropy θ_q would be equal to the data density h_d, i.e.,

$$\theta_{q0} = -\sum_{i=1}^{\kappa} \frac{1}{\kappa} \log_2 \kappa = \eta = h_d.$$

However, neighboring samples ordinarily are not independent. As shown in Fig. 4(a), significant correlation exists out to approximately three neighboring samples (in all directions) if the mean spatial detail $\mu = 1$, and out to about 10 neighboring samples if $\mu = 3$. Blurring in the image-gathering process will further increase the correlation among neighboring samples for the fine spatial detail. In practice, it is common to consider only the nearest samples as depicted in Fig. 10. If we consider only the past nearest neighbor [see Fig. 10(b)], then the

corresponding entropy θ_{q1} is defined by

$$\theta_{q1} = -\frac{1}{2}\sum_{i=1}^{\kappa}\sum_{j=1}^{\kappa} p(x_i, x_j)\log_2(x_i, x_j). \tag{12}$$



(a) Independent points, θ_{p0} (b) Past nearest neighbor, one dimensional, θ_{p1} (c) Past nearest neighbors, two dimensional, θ_{p3} (d) Nearest neighbors, two dimensional, θ_{p8}

Figure 10. Samples used to estimate the entropy of the digital data $s_q(x, y)$.

The entropies θ_{q3} and θ_{q8} that include, respectively, the past three nearest samples [Fig. 10(c)] and the 8 nearest samples [Fig. 10(d)] are defined similarly. Hence, as we include an in creasing number of neighbors, we approach the entropy θ_q defined by Eq. (10).

Figure 11 illustrates the variation of the entropies θ_{q0}, θ_{q1}, θ_{q2}, and θ_{q3} with the mean spatial detail μ. The zeroth order entropy θ_{q0} does not account for any of the correlation that exists among the neighboring samples. It depends solely on the variance of the incident radiance field and on the effects of the image-gathering process (including quantization). The magnitude of the higher order entropies decreases as more of the correlation among the neighboring samples is accounted for. The independence of the higher entropies from the properties of the radiance field and image-gathering system are probably limited to the informationally optimized designs considered in this paper.

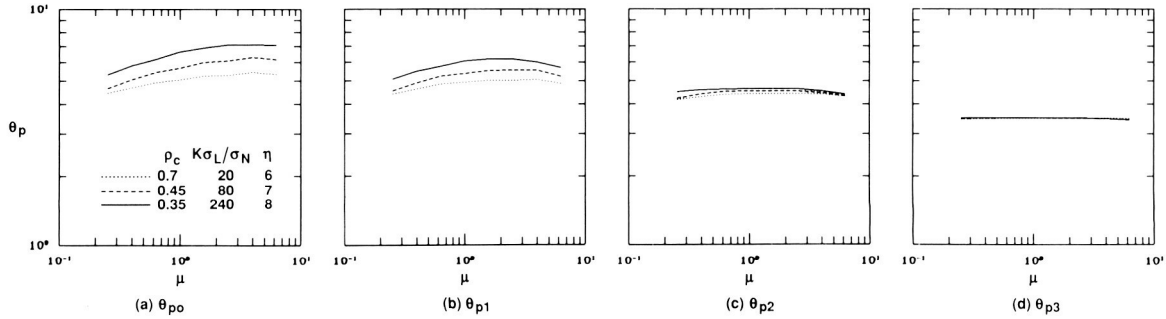


Figure 11. Estimates of the entropy θ_q versus the mean spatial detail μ. The image-gathering process is characterized by the optical design parameter ρ_c, the SNR $K\sigma_L/\sigma_N$, and the encoding level η.

Another limitation of these results is that the absolute magnitude of the entropies shown in Fig. 11 cannot be compared strictly to the information density h_q computed by Eq. (9) and shown in Figs. 8 and 9. The reason is that the entropies are obtained from a single target, such as shown in Fig. 3, in which the rectangles have a fixed orientation, whereas the information densities are computed for a circularly symmetric Wiener spectrum derived with the assumption that the orientation of these rectangles is random with a uniform distribution. Hence, the estimates of information efficiency given below will err on the high side. The reason is that the entropy would be higher for the actual radiance field with random edge orientations than for the simulated radiance field with a fixed edge orientation.

B. Image Coding

Figure 12 illustrates the lossless data compression method that we use to assess the effect of compression on the information efficiency of the transmitted data. The purpose of the compression is to translate the string of quantized data $s_q(x, y)$ into an encoded string of data $\Delta s_e(x, y)$ that is (usually) a compressed version of $s_q(x, y)$.

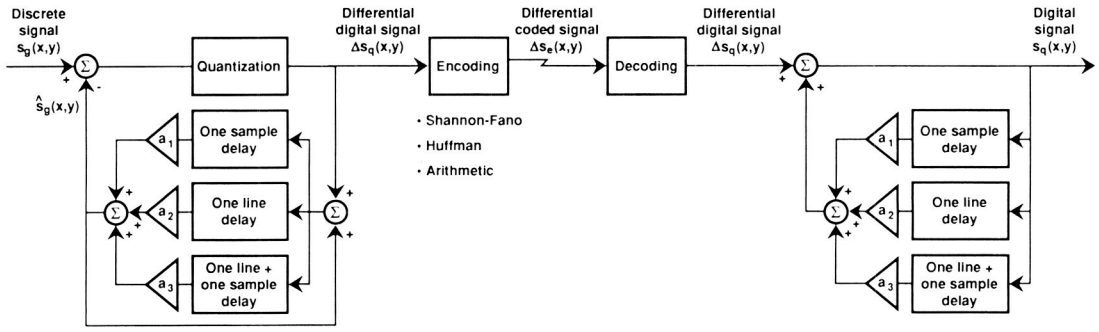


Figure 12. Model of lossless predictive compression that uses the past three nearest neighboring samples [as depicted in Fig. 10(c)].

The predictive compression reduces the redundancy, or correlation, of the string of digital data $s_q(x, y)$ prior to the encoding. The system shown in Fig. 12 is often referred to as a third-order predictor because it uses the values of the three past nearest neighbors to predict the value that is about to be read out. We select the weighting values a_1, a_2, and a_3 so that the linear mean-square error estimation $E\left\{[s_g(x, y) - \hat{s}_g(x, y)]^2\right\}$ is minimized. This minimization, which is referred to as best linear estimate, is commonly favored because of its mathematical tractability even though some improvement in performance can often be gained when nonlinear functions are used to form the estimate.[9] For the wide-sense stationary input radiance field, the correlation between neighboring samples is independent of location and the three values a_1, a_2, and a_3 can be computed as follows. Let $R(m, n)$ be the (normalized) correlation between the samples located at $(i - m, j - n)$ and (i, j) [see Fig. 10(c)], then the desired predictor weighting values are given by the following three simultaneous equations:[9]

$$\begin{bmatrix} R(0,1) \\ R(1,1) \\ R(1,0) \end{bmatrix} = \begin{bmatrix} R(0,0) & R(1,0) & R(1,1) \\ R(1,0) & R(0,0) & R(0,1) \\ R(1,1) & R(0,1) & R(0,0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix},$$

where $R(0, 0) = 1$.

Figure 13 gives the correlation values $R(m, n)$ and the predictor weighting values a_1, a_2, and a_3 for three radiance fields and three image-gathering systems. The radiance fields are characterized by the mean spatial detail μ, and the image-gathering systems are characterized by the optical-design parameter ρ_c, the SNR $K\sigma_L/\sigma_N$, and the number of encoding levels η. The behavior of the correlation and weighting values appeals intuitively. The correlation between neighboring samples increases both as the mean spatial detail becomes larger (i.e., as μ increases) and as it becomes more blurred (i.e., as ρ_c decreases). In the limit, the sum of the weighting values (i.e., $a_1 + a_2 + a_3$) approaches unity, which suggests that the new sample will be similar in value to the neighboring ones with increasing probability. It also is interesting to observe that the correlation $R(0, 1)$ and $R(1, 0)$ of immediately neighboring samples lies between 0.84 and 0.92 when the mean spatial detail is $\mu = 3$. This result turns out to be in close agreement with the observation made by Gonzalez and Wintz[6] (pg. 298) that, in practice, this correlation typically lies between 0.85 and 0.95 for properly sampled images.

The image coding achieves further compression by transmitting the more probable symbols in fewer bits than the less probable ones. The *Huffman* code[5,19] that we use is derived by

successively merging the two least probable samples of $\Delta s_q(x, y)$ into a new sample which is assigned a probability equal to the sum of the former two probabilities. This process is continued until exhaustion. The result of this process is arranged as a tree that is used to determine the code words for the quantized data.

Figure 14 characterizes the effects of the data compression. The compression h_d/h_e is given by the ratio of data density $h_d = \eta$ without coding to data density h_e with coding. This compression does not vary significantly with either the mean spatial detail or the design of the image-gathering system. The compression remains within the range of 1.6 to 1.9, and thus approaches the factor of 2 that is often given for lossless data compression. However, the information efficiency h_q/h_e of the encoded data depends significantly on the image-gathering system design. These results suggest the following generalization:

(4) The upper limit of the information efficiency achieved with lossless data compression increases from $\sim$ 0.5 to 0.9 bif/bit as the information density of the encoded data increases from $\sim$ 2 to 4 bifs. Thus, high information density is transmitted more efficiently than low information density.

This generalization intuitively is appealing since the encoded data contain less image-gathering degradation (e.g., aliasing and noise) when the information density is high rather than low.

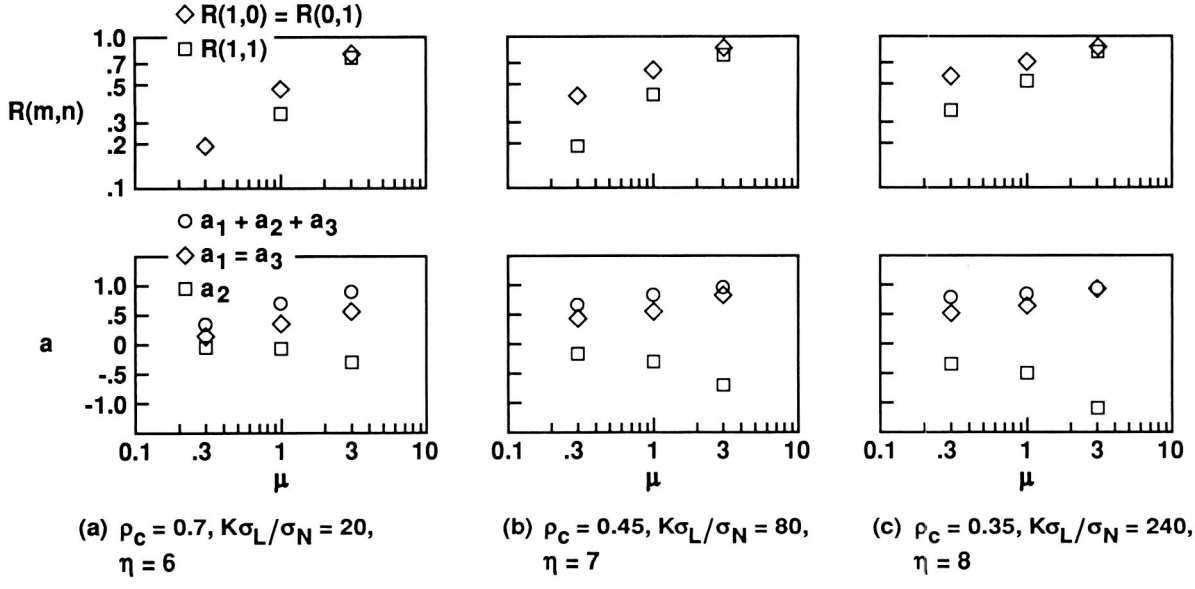


Figure 13. Characteristics of the lossless predictive compressor as a function of the mean spatial detail μ. Shown are the (normalized) correlation $R(m, n)$ between the neighboring samples and the predictor weighting values a.

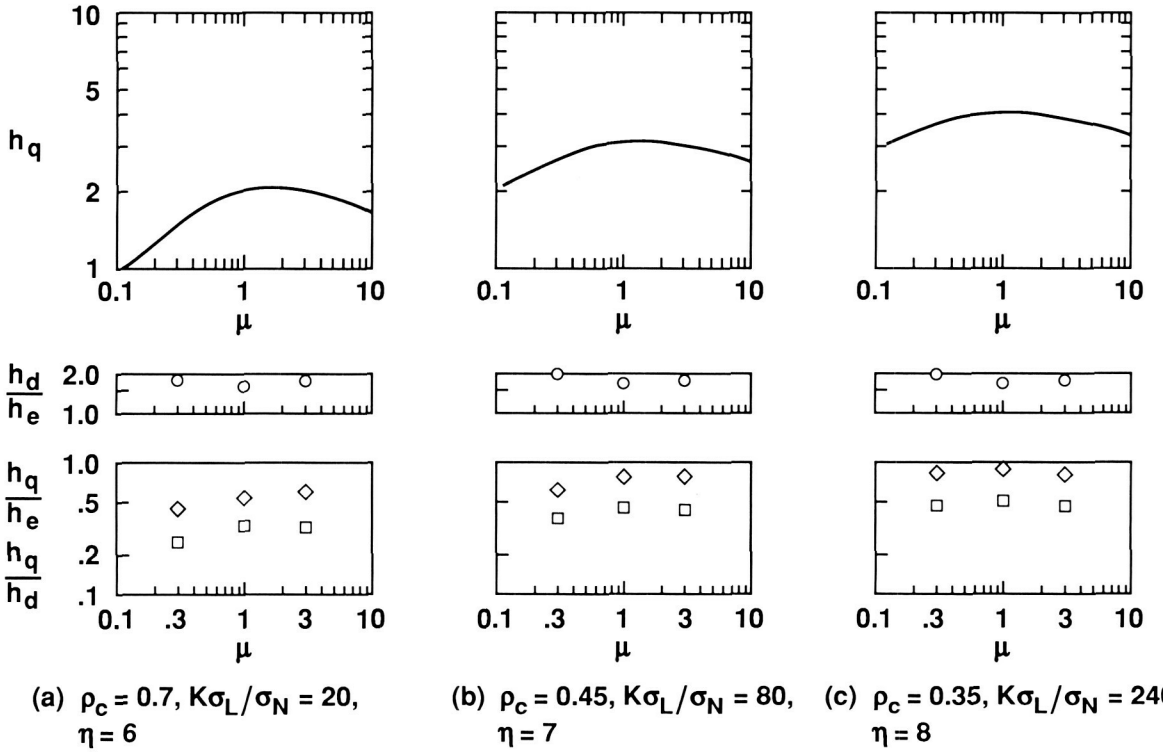


Figure 14. Characteristic of the encoded signal as a function of the mean spatial detail μ. Shown are the information density h_q, the data compression h_d/h_e, and the information efficiency h_q/h_d and h_q/h_e before and after compression, respectively.

5. IMAGE RESTORATION

A. Information and Fidelity

The data-processing algorithm that maximizes the fidelity of the restored image is given by the unconstrained Wiener filter[1-4]

$$\hat{\Psi}(v,w) = \frac{\hat{\Phi}'_L(v,w)\hat{\tau}_g^*(v,w)}{\hat{\Phi}'_L(v,w)|\hat{\tau}_g(v,w)|^2 * \underline{\underline{|||}}(v,w) + \left(\frac{K\sigma_L}{\sigma_N}\right)^{-2} + \kappa^{-2}}. \tag{13}$$

If the radiance-field spectrum $\hat{\Phi}'_L(v,w)$, the image-gathering response $\hat{\tau}_g(v,w)$, and the SNR $K\sigma_L/\sigma_N$ are exactly accounted for in $\hat{\Psi}(v,w)$, then the image fidelity f reaches its maximum realizable value f_m given by[1-4]

$$f_m = \int\int_{-\infty}^{\infty} \hat{\Phi}_L(v,w)\hat{\tau}_g(v,w)\hat{\Psi}(v,w)dv\ dw. \tag{14}$$

It also is possible, then, to express the Wiener filter $\hat{\Psi}(v,w)$ and the fidelity f_m as a function of the spectral information density $\hat{h}_q(v,w)$ as follows:[1,2]

$$\hat{\Psi}(v,w) = \frac{1}{\hat{\tau}_g(v,w)}\left[1 - 2^{-\hat{h}_q(v,w)}\right]$$

and

$$f_m = \iint_{-\infty}^{\infty} \hat{\Phi}'_L(v,w) \left[1 - 2^{-\hat{h}_q(v,w)}\right] dv\, d_w,$$

where $\hat{h}_q(v,w)$ is the integrand of Eq. (9) given by

$$\hat{h}_q(v,w) = \log_2 \left[1 + \frac{\hat{\Phi}'_L(v,w)|\hat{\tau}^*_g(v,w)|^2}{\hat{\Phi}'_L(v,w)|\hat{\tau}_g(v,w)|^2 * \widehat{\underset{\neq 0,0}{|||}}(v,w) + \left(\frac{K\sigma_L}{\sigma_N}\right)^{-2} + \kappa^{-2}}\right].$$

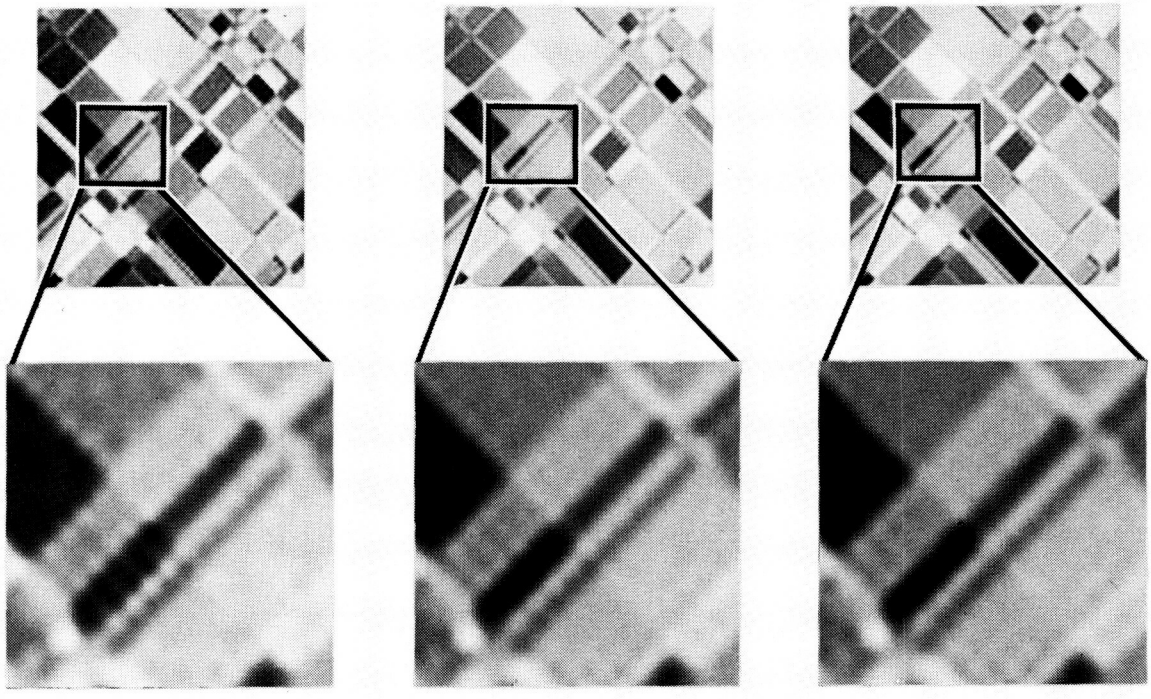
These relationships show that our ability to restore images (restorability) is solely limited by the term $2^{-\hat{h}_q(v,w)}$.

Since the restorability depends on the *spectral* information density $\hat{h}_q(v,w)$ rather than on the *total* information density h_q, it is not possible to directly ascertain whether increases in the information density h_q will always increase the restorability. Nevertheless, it seems reasonable to expect that the restorability of images ordinarily will be correlated positively to the available information density.

B. Fidelity-Maximized Restorations

Figure 15 presents fidelity-maximized images for the three informationally optimized designs characterized in Fig. 14. The change in the visual quality that occurs with increasing information density manifests itself mainly as an increase in the resolution, contrast, and clarity. Noise and aliasing artifacts disappear almost entirely as the highest available information density is approached. However, ringing near sharp edges now becomes a major visual defect. This ringing (Gibbs phenomenon) occurs because of the steep roll-off in the Wiener filter. As Schreiber[5] (pg. 92) summarizes: "... it is impossible to have maximum sharpness with neither aliasing nor ringing, all at the same time. Since the latter is least acceptable, some aliasing and loss of sharpness must be accepted by using a more gradual roll-off in the filter." These results suggest the following important generalization:

(5) Increases in the information density in the fidelity-maximized images is perceived mostly as a decrease in image-gathering degradations (i.e., aliasing artifacts and noise). However, defects (i.e., ringing) introduced by the image-restoration process become more apparent and may limit the amount of information that is useful for image restoration.



(a) $\rho_c = 0.7$, $K\sigma_L/\sigma_N = 20$, $\eta = 6$; $h_q = 2$, $h_e = 3.2$, $h_q/h_e = 0.62$

(b) $\rho_c = 0.45$, $K\sigma_L/\sigma_N = 80$, $\eta = 7$; $h_q = 3.1$, $h_e = 3.9$, $h_q/h_e = 0.79$

(c) $\rho_c = 0.35$, $K\sigma_L/\sigma_N = 240$, $\eta = 8$; $h_q = 3.9$, $h_e = 4.6$, $h_q/h_e = 0.84$

Figure 15. Images restored with the Wiener filter for three informationally optimized image-gathering systems. The systems are characterized by the optical design parameter ρ_c, and SNR $K\sigma_L/\sigma_N$, and the encoding level η. The transmitted data are characterized by information density h_q, data density h_e, and information efficiency h_q/h_e. The mean spatial detail $\mu = 3$.

Figure 16 illustrates the dependence of the image fidelity f_m on the mean spatial detail μ for the three image-gathering systems characterized by Figs. 14 and 15. As can be seen, the image fidelity depends almost solely on the characteristics of the target, even though the resolution, sharpness, and clarity of the images restored for maximum fidelity depend perceptibly on the available information density. This result suggests the following generalization:

(6) Image fidelity is not a suitable criterion for assessing the performance of image gathering and coding. Not only is it insensitive to the visual flaws of the fidelity-maximized images but also to the improvements that are gained in the visual quality of these restorations with increasing information density.

C. Restorations for Visual Quality

The final step in restoring images for maximum visual quality still must be based on perceptual rather than mathematical considerations. For this reason it is necessary to introduce some ad hoc modification of the Wiener filter to control adaptively the trade-off between the enhancement of spatial detail and the suppression of visual defects. Unfortunately, these adaptive controls reduce the quantitative connection between optimum filtering and informationally optimized image gathering that we have tried to maintain so far. This seems unavoidable as long as visual quality cannot be assessed by some figure of merit.

The goals of the ad hoc modification of the Wiener filter are to reduce the ringing at sharp edges and to enhance the visibility of the fine detail and of the boundaries between areas much

larger than those that are barely perceived. To provide these adaptive controls, McCormick et al.[4] introduced the Wiener-Gaussian enhancement (WIGE) filter given by

$$\widehat{\Psi}_{ie}(v,w) = \widehat{\Psi}(v,w)\left\{\exp\left[-2(\pi\sigma_i\rho)^2\right] + \zeta(w\pi\rho)^2\exp\left[-2(\pi\sigma_e\rho)^2\right]\right\}, \tag{15}$$

where ζ is the enhancement parameter that controls the relative amount of the synthetic-high filtered frequency components in the restored image. The standard deviation σ_i controls the smoothing of the low-pass filtered image, and the standard deviation σ_e controls the smoothing associated with the edge enhancement. Nevertheless, a trade-off remains between the enhancement of fine detail and sharp edges and the suppression of ringing.

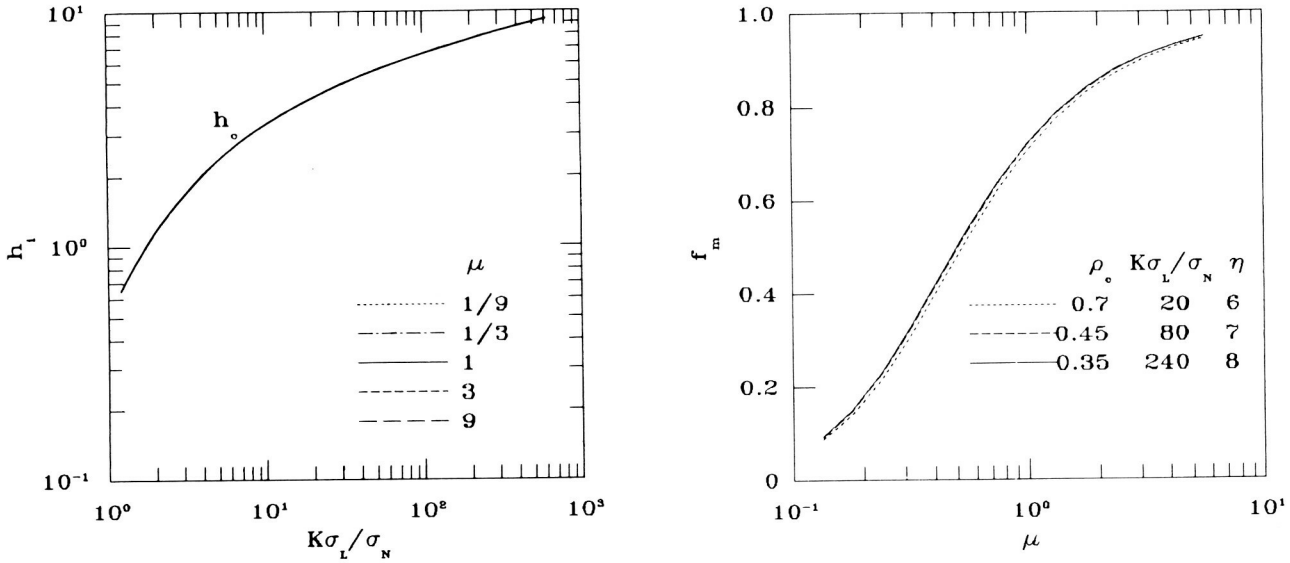
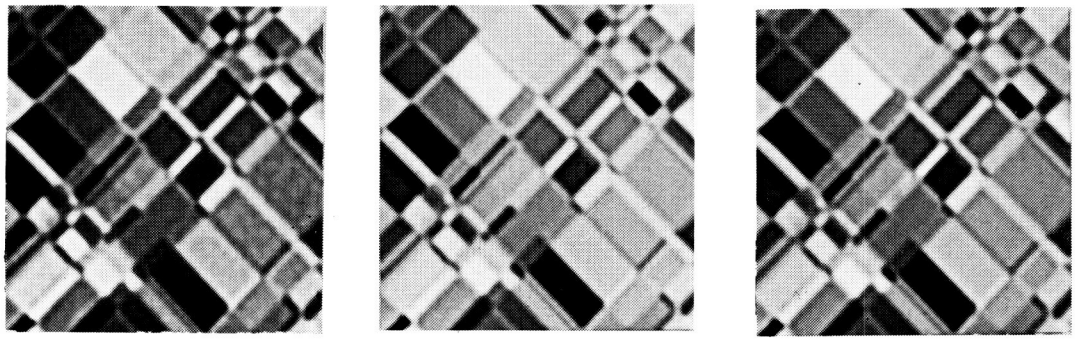


Figure 16. Image fidelity f_m versus the mean spatial detail μ. The image-gathering system is characterized by the optical-design parameter ρ_c, the SNR $K\sigma_L/\sigma_N$, and the encoding level η.

The preferred values for the WIGE parameters depend on the target, the design of the image-gathering system, and the objectives of the observer. For example, if the image-gathering system is informationally optimized for high SNR's and the target is the random radiance field used here as an example, then the trade-off between the suppression of ringing and the loss of sharpness in the fidelity-maximized images shown in Fig. 15 is reasonably well resolved with $\sigma_i = 0.4$ and $\sigma_i = 0.8$. The contrast of the fine detail and sharp edges become enhanced increasingly as ζ is increased. However, this enhancement is achieved only at the cost of general visual quality as well as fidelity. Depending on the objectives of the observer, the preferred value for ζ ranges typically from 0.2 to 0.8.[4]

Figure 17 presents images restored with the WIGE filter. As above, for the fidelity-maximized images shown in Fig. 15, the improvement in image quality with increasing information density is perceived mostly as an increase in clarity. Noise and aliasing artifacts disappear almost entirely as the highest available information density is approached. Moreover, ringing near sharp edges now has been suppressed effectively but at some cost in resolution and sharpness. A small overshoot still occurs at the boundaries between areas much larger than those that are barely perceived. This overshoot enhances the visibility of the boundaries and therefore is often preferred; however, it can be suppressed by reducing the enhancement parameter ζ. These results suggest the following generalization:

(7) The visual quality that can be attained with interactive image restoration improves perceptibly as the information density increases to $\sim$ 3 bifs. However, the perceptual improvements that can be gained with further increases in information density are very subtle.



(a) $\rho_c = 0.7$, $K\sigma_L/\sigma_N = 20$, $\eta = 6$; $h_q = 2$, $h_e = 3.2$, $h_q/h_e = 0.62$

(b) $\rho_c = 0.45$, $K\sigma_L/\sigma_N = 80$, $\eta = 7$; $h_q = 3.1$, $h_e = 3.9$, $h_q/h_e = 0.79$

(c) $\rho_c = 0.35$, $K\sigma_L/\sigma_N = 240$, $\eta = 8$; $h_q = 3.9$, $h_e = 4.6$, $h_q/h_e = 0.84$

Figure 17. Images restored with the WIGE filter for three informationally optimized image-gathering systems. The conditions are the same as in Fig. 15.

The relationship between the available information density and the visual quality of the interactively restored images can be expected to depend significantly on the properties of the target. Hence, this relationship still must be assessed for a variety of different targets and enhancements.

6. CONCLUDING REMARKS

The goal of data compression, as it is traditionally stated, is to reduce, as much as possible, the number of bits necessary to reconstruct a faithful duplicate of the original picture. The effects of data compression are assessed qualitatively by visually comparing the duplicate to the original picture. This assessment ignores entirely the degradations that image gathering and reconstruction introduce into the original picture. It also ignores the potential capabilities of digital processing to reduce the visibility of the degradations caused by image gathering and reconstruction and to enhance various features for close scrutiny.

Our point of view is closer to the one that Schreiber[10] expressed with the question: "For a given channel, what relationship between the original scene and the transmitted signal produces the 'best' picture?" Clearly, this relationship depends on the combined performance of image gathering and coding. However, our constraints, aside from the communication bandwidth, differ from those of Schreiber. Schreiber was concerned mostly with telephotography and televison. These applications are constrained, for commercial reasons, mostly by the cost of the image display. By contrast, we are concerned with space activities and planetary exploration. These applications, in turn, are constrained mostly by the size, weight, and power limitations imposed on the spacecraft instrumentation. The complexity of the digital processing required to restore images and enhance features is ordinarily not a critical constraint. The latter situation also arises frequently in military reconnaissance and medical diagnosis.

Thus, the goal of this paper has been to develop a method for assessing the combined performance of image gathering and coding in terms of the information density and efficiency of the transmitted data. This method is based on earlier findings[1,2] that informationally optimized image gathering maximizes the fidelity of the images restored by the Wiener filter, provided that this filter accounts correctly for the image-gathering degradations. However, an important obstacle remains: the fidelity-maximized images exhibit some visual defects such as ringing, aliasing artifacts, and noise.[3] Therefore, in practice, it is often desirable to reduce these defects with interactive processing.[4]

The method for optimizing the end-to-end performance of image gathering and coding for interactive restoration that is suggested in this paper is (1) to assess how much information

is required to restore and enhance images with sufficiently high visual quality for a particular application, and (2) to assess how this information can be acquired and encoded most efficiently. The preliminary results that we have presented are limited to a single, artificial target. However, these results intuitively are attractive and consistent with practical experience.

REFERENCES

1. F. O. Huck, C. L. Fales, N. Halyo, R. W. Samms, and K. Stacy, "Image gathering and processing: Information and fidelity," J. Opt. Soc. Am. A2, pp. 1644-1666 (1985).
2. F. O. Huck, C. L. Fales, J. A. McCormick, and S. K. Park, "Image-gathering system design for information and fidelity," J. Opt. Soc. Am. A5, pp. 285-299, March 1988.
3. C. L. Fales, F. O. Huck, J. A. McCormick, and S. K. Park, "Wiener restoration of sampled image data: End-to-end analysis," J. Opt. Soc. Am. A5, 300-314 (1988).
4. J. A. McCormick, R. Alter-Gartenberg, and F. O. Huck, "Image gathering and restoration: Information and visual quality," J. Opt. Soc. Am. A6, 987-1005 (1989).
5. H. C. Andrews and B. R. Hunt, Digital Image Restoration (Prentice-Hall, Englewood Cliffs, N. J., 1977).
6. R. C. Gonzalez and P. Wintz, Digital Image Processing (Addison-Wesley, Reading, Massachusetts, 1977).
7. W. K. Pratt, Digital Image Processing (Wiley, New York, 1978).
8. T. S. Huang, ed., Picture Processing and Digital Filtering (Springer-Verlag, Berlin, 1979).
9. A. Rosenfeld and A. C. Kak, Digital Picture Processing (Academic, New York, 1982).
10. W. F. Schreiber, Fundamentals of Electronic Imaging Systems (Springer-Verlag, Berlin, 1986).
11. C. E. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J. 27, 379-423, and 28, 623-656 (1948); C. E. Shannon and W. Weaver, The Mathematical Theory of Communication (U. Illinois Press, Urbana, 1964).
12. P. B. Fellgett and E. H. Linfoot, "On the assessment of optical images," Philos. Trans. R. Soc. London 247, 369-407 (1955).
13. J. W. Modestino and R. W. Fries, "Edge detection in noisy images using recursive digital filtering," Comput. Vision Graphics Image Process. 6, 409-433 (1977).
14. Y. Itakura, S. Tsutsumi, and T. Takagi, "Statistical proper ties of the background noise for the atmospheric windows in the intermediate infrared region," Infrared Phys. 14, 17-29 (1974).
15. M. Kass and J. Hughes, "A stochastic image model for AI," IEEE Int. Conf. on Systems, Man, and Cybernetics, 369-372 (1983).
16. C. B. Johnson, "A method for characterizing electro-optical device modulation transfer functions," Photogr. Sci. Eng. 15, 413-415 (1970).
17. A. I. Khinchin, Mathematical Foundations of Information Theory (Dover, New York, 1957).
18. R. C. Jones, "Information capacity of a beam of light," J. Opt. Soc. Am. 52, 493 (1962).
19. D. A. Huffman, "A method for the construction of minimum redundancy codes," Proc. IRE 40, No. 9, 1098-1101 (September 1952).